METHOD AND APPARATUS FOR SPEECH RECOGNITION INCORPORATING LOCATION INFORMATION

BACKGROUND OF THE INVENTION

The present invention relates generally to the field of speech recognition and, more particularly, to a method for improving the likelihood that voice input to a speech recognizer is correctly recognized.

Speech recognizers are used in a variety of applications. One exemplary application is voice-dialing in communication devices such as cellular phones. Voice-dialing allows a user to speak a label such as "Robert" to dial a corresponding telephone number, saving the user the time and effort to key in the telephone number one digit at a time. In typical voice-dialing implementations, the communication device accesses a stored directory comprising patterns representing spoken labels and their corresponding telephone numbers. When the user says a particular label, the voice input is processed to produce a pattern representing the voice input. This pattern is compared to each of the stored patterns in the directory, and a confidence score is returned for each pattern indicating the likelihood that the voice input corresponds to that particular pattern

and, hence, the associated label. If the voice input is not recognized correctly, an incorrect number may be dialed, embarrassing or inconveniencing the user. Therefore, there is a need for methods of increasing the likelihood that the correct number will be dialed.

SUMMARY OF THE INVENTION

A method and apparatus are provided for speech recognition incorporating location information.

The method comprises receiving voice input from a user, associating a confidence measure with each of a set of stored patterns according to how closely the stored pattern matches the voice input, acquiring location information, and modifying one or more of the confidence measures based upon the location information.

The apparatus comprises a speech recognition circuit that provides confidence measures, circuitry for acquiring location information, and logic for modifying the confidence measures based on the location information.

In one embodiment, the invention improves the performance of voice-dialing in a mobile communication device.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1a is a functional block diagram of an environment of the invention.

Fig. 1b is a functional block diagram of a mobile communication device in which voice-dialing may be implemented.

Fig. 2 is a flowchart illustrating the general operation of the invention in a voice-dialing application.

Fig. 3 is a flowchart illustrating the operation of an aspect of the invention.

Fig. 4 is a flowchart illustrating the operation of another aspect of the invention.

Fig. 5 is an exemplary table illustrating a record of previous telephone calls in accordance with yet another aspect of the invention.

Fig. 6 is a flowchart illustrating the operation of yet another aspect of the invention that uses the kind of information illustrated in Fig. 5.

Fig. 7 is a flowchart illustrating the operation of a further aspect of the invention.

Fig. 8 is an exemplary table illustrating stored patterns and their corresponding confidence measures before being modified.

Fig. 9 is an exemplary table illustrating stored patterns and their corresponding confidence measures after modification by the invention.

DETAILED DESCRIPTION OF THE INVENTION

The invention operates within an environment 100 such as that shown in Fig. 1a. A speech recognition circuit 110 receives voice input 115, processes it to produce a

pattern representing voice input 115, compares the pattern to each of a set of patterns stored in memory 120, produces a confidence measure for each stored pattern, and stores the confidence measures in memory 120. A confidence measure is a numeric indication of the degree to which the pattern representing voice input 115 matches the associated stored pattern. Those skilled in the art will recognize that it is convenient to express a confidence measure as a percentage or a number between 0 and 1, inclusive. A circuit for acquiring location information 130 outputs location information 140 to logic 150. The circuit 130 for acquiring location information may comprise a Global Positioning System (GPS) receiver. Circuit 130 may alternatively use techniques such as terrestrial cellular positioning or coarse positioning using the base station identification codes of one or more base stations, all of which are known in the applicable art. In some applications, circuit 130 may use exchange codes, area codes, or country codes included in telephone numbers as proxy for more precise location information. Logic 150 modifies the confidence measures in memory 120 in accordance with location information 140. The modified confidence measures in memory 120 may then be used for decision making by a suitable speech recognition application 160.

For most of the remainder of this description, the invention is described within the context of an exemplary application, namely voice-dialing in a mobile communication device such as a cellular telephone, Personal Communication System (PCS) telephone, satellite

telephone, cordless telephone, pager, personal digital assistant, or communicator. In the voice-dialing application to be described, the environment 100, therefore, is part of a mobile communication device 165, for which a functional block diagram is shown in Fig. 1b. Mobile communication device 165 comprises a microphone 170 and microprocessor 175, which communicates over data bus 180 with circuit 130, speech recognition circuit 110, input device 183, memory 185, display 190, and RF transceiver 195. RF transceiver 195 is electrically connected to antenna 198. Although the invention is described primarily in terms of voice-dialing in mobile communication device 165, the invention applies to any situation in which speech recognition may be aided by location information and is not limited to the voice-dialing application.

Fig. 2 is a flowchart of the general operation of the invention in a voice-dialing application within a mobile communication device 165. At 210, microprocessor 175 determines whether the voice-dialing function has been activated. For example, the user may activate voice-dialing via input device 183, which may comprise a keypad. If voice-dialing has not been activated, its next use is awaited at 215. Once voice-dialing has been activated, voice input 115 from a user is accepted at 220. Voice input 115 comprises a spoken label that the user has previously associated with a telephone number to be dialed. The label may comprise a word, name, phrase, nickname, or other utterance appropriate for the type of speech recognition technology employed. After receiving

voice input 115, speech recognition circuit 110, according to methods known in the applicable art, processes voice input 115 to produce a pattern representing voice input 115. Speech recognition circuit 110 compares the pattern representing voice input 115 with each of a set of stored patterns, each of which in turn corresponds to a stored telephone number, and returns a confidence measure for each pattern at 225. The confidence measures are optionally ranked from greatest to smallest at 230. Although those skilled in the art will recognize that ranking the confidence measures at 230 is not essential, doing so simplifies the notation in the description of the next few steps and has been included in Fig. 2 for that purpose. Therefore, without loss of generality, let the ranked confidence measures, in descending order, be denoted as M_1 through M_N, respectively. At 240, the difference between the greatest and next-to-greatest confidence measures, M_1 and M_2, is computed. If the difference exceeds a predetermined threshold T, the decision is assumed to be unambiguous, and the party corresponding to confidence measure M_1 is dialed at 245. If the difference at 240 does not exceed T, an ambiguous situation occurs, and circuit 130 attempts to determine the current location of environment 100 at 250. If the current location of mobile communication device 165 is not available at 260, the party corresponding to confidence measure M_1 is dialed. Otherwise, location information 140 is fed to logic 150 at 270. According to various aspects of the invention to be described more fully in subsequent

portions of this description, the confidence measures are modified at 275 in accordance with location information 140 to produce the confidence measures denoted M'_1 through M'_N. At 280, the party with the largest modified confidence measure is dialed.

In one variation of the general operation depicted in Fig. 2, location information 140 is acquired unconditionally, and the confidence measures are modified based on location information 140 regardless of whether or not an ambiguous situation occurs. Therefore, in such a variation, it is not necessary to compute the difference between the greatest and next-to-greatest confidence measures, M_1 and M_2, before carrying out steps 270 and 275. As in the foregoing explanation of Fig. 2, ranking the confidence measures at 230 is not essential.

The flowchart of Fig. 3 shows the operation of one aspect of the invention, more particularly an implementation of steps 270 and 275 in Fig. 2. This particular aspect employs distance between the called party and the current location of mobile communication device 165 as a criterion for modifying the confidence measures. In this aspect of the invention, the current location of mobile communication device 165 is assumed known from circuit 130. One method of determining the distance between mobile communication device 165 and a called party is to consult a look-up table mapping sub-parts of telephone numbers such as exchange codes, area codes, or country codes to their corresponding geographical regions. Once an approximate location corresponding to a stored telephone number is known, an

approximate distance may be computed between mobile communication device 165 and the stored telephone number using a second look-up table containing distances between various geographical regions such as cities, towns, or metropolitan areas. One efficient method to implement this aspect of the invention is to store the approximate location corresponding to a stored telephone number in the same directory or database as the stored telephone numbers and their corresponding patterns. In Fig. 3, a counter is initialized at 310, and the distance between the current location of mobile communication device 165 and the telephone number associated with the first stored pattern is computed at 320. At 330, the distance is compared to a predetermined threshold D_s. For example, D_s might be 15 miles. If the distance is less than D_s, the confidence measure associated with the first stored pattern is increased at 340. For example, the confidence measure might be increased by 10 per cent when the distance is less than D_s and not increased when the distance is greater than or equal to D_s. Optionally, the confidence measure is instead decreased if the distance is greater than or equal to D_s and left unmodified if the distance is less than D_s. In yet another variation, the confidence measure is increased if the distance is less than D_s and decreased if the distance is greater than or equal to D_s. All of the foregoing variations of the modification at 340 have the effect of favoring, in the final selection of the number to be dialed at 280, those stored telephone numbers that are close to the current location of mobile communication device 165 over those

that are more distant. At 350, the counter is incremented. Steps 320 through 350 are then repeated until the test at 360 determines that all stored patterns have been considered. At 370, control is returned to 280 in Fig. 2.

Fig. 4 illustrates the operation of another aspect of the invention, more particularly another implementation of steps 270 and 275 in Fig. 2. This aspect employs sub-parts of telephone numbers such as area codes, exchange codes, or country codes as a proxy for more precise location information in a different manner from the aspect shown in Fig 3. Fig. 4 differs from Fig. 3 at 410, where a sub-part of the stored telephone number corresponding to each stored pattern is tested using a look-up table such as the one described in connection with Fig.3 to determine whether or not the stored telephone number corresponds to a geographical region within which mobile communication device 165 lies. In this aspect of the invention, no distances are computed. The current location of mobile communication device 165 is assumed to be known to within the resolution of the applicable geographical region corresponding to the sub-part of the stored telephone number being tested at 410. For example, the current location of mobile communication device 165 may be known to be within area code 919 based on terrestrial cellular positioning data obtained from circuit 130. In general, the geographical region may be that of a single area code, several area codes (a metropolitan area), the region served by an exchange code, or a country, depending on which sub-part is

examined. Those skilled in the art will recognize that which sub-part to examine is a design choice that depends, at least in part, on the degree of granularity that can be tolerated in the geographical regions upon which the test at 410 is based. If a match occurs at 410, the associated confidence measure is modified as described in connection with Fig. 3. This aspect of the invention is intended to aid a user who travels frequently to one or more geographical regions. For example, a person traveling from Raleigh, NC to New York City and walking along a street within area code 212 would likely prefer that a telephone number with a label "Bob" located within New York City area code 646 be favored over another number with the same label located in Portland, OR when an ambiguous situation occurs. Since some large metropolitan areas such as New York City have multiple area codes, the test at 410 uses a look-up table mapping area codes to cities to determine that the "Bob" who lives in area code 646 and the user are in the same metropolitan area.

Fig. 5 is an exemplary table illustrating a record of previous telephone calls in accordance with yet another aspect of the invention, the operation of which is illustrated in Fig. 6. In Fig. 5, the left column 510 contains dates on which the out-going calls were placed using mobile communication device 165. The phone numbers dialed are contained in the center column 520. The right column 530 contains named locations such as "Home," "Office," or "Vacation Home" from which the calls were placed. The user defines the named locations by

associating a name with geographical coordinates or other location information obtained from circuit 130. Mapping geographical coordinates to user-defined place names is known in the art and is described in more detail in U.S. Patent 6,078,826, which is also assigned to Ericsson Inc.

Fig. 6 illustrates the operation of yet another aspect of the invention, more particularly an implementation of steps 270 and 275 in Fig. 2. When an ambiguous situation occurs, this aspect uses a call history record including correlated location information such as that shown in Fig. 5 to favor stored telephone numbers that have been called frequently from the current location of mobile communication device 165. Fig. 6 is much like Fig. 3, except that the relative frequency with which each stored telephone number has been called previously from the current location of mobile communication device 165 is computed at 610, and the computed relative frequency is compared to a threshold F at 620. The confidence measure is then modified in a manner analogous to that described in connection with Fig. 3, the difference being that the threshold is now one of frequency instead of distance. This aspect of the invention aids voice-dialing based on the assumption that many users tend to call a given stored telephone number from the same location much of the time. For example, a worker may call his spouse from the office frequently. Suppose the spouse's name is "Peggy" and that the worker has a head cold when he attempts to voice-dial from the office on a particular occasion, resulting in an ambiguous situation in which both "Peggy" and the

worker's out-of-state sister "Penny" are assigned approximately the same confidence measure by the speech recognition circuit 110. The worker never calls his out-of-state sister from the office, so the aspect of the invention shown in Fig. 6 favors "Peggy" over "Penny" by increasing the confidence measure associated with "Peggy" relative to "Penny."

Fig. 7 illustrates the operation of a further aspect of the invention, more particularly an implementation of steps 270 and 275 in Fig. 2. In this aspect, which is related to the aspect shown in Fig. 6, the user defines a list of most frequently called telephone numbers for each of a set of named locations defined by the user. At 710, the stored telephone numbers corresponding to the stored patterns are compared to the stored list of most frequently called telephone numbers associated with the current location of mobile communication device 165. If the test at 710 determines that a stored telephone number matches one of the numbers on the list, the corresponding confidence measure is modified as described in connection with Fig. 3. One advantage of this approach is that the user may intervene to change the behavior of the invention when circumstances change suddenly, at the expense of some extra time spent in configuration. Further, this approach does not require the computation of relative frequencies, nor does it require a record of previous calls.

All of the foregoing aspects of the invention have the effect of modifying the confidence measures returned by speech recognition circuit 110 based on location

information 140. An example of how this works in practice and how the dialing decision at 280 is affected is illustrated in Figs. 8 and 9, in which a user has attempted to voice-dial his brother "Irvin" from the user's home. Fig. 8 shows exemplary labels 810 and their associated confidence measures 820 before being modified based on location information 140. The situation illustrated is ambiguous because the difference between the confidence measure associated with "Irvin" and a co-worker named "Mervyn" is slight. In Fig. 9, after the confidence measures 820 have been modified based on location information 140, the confidence measures 910 associated with "Irvin" and a golf buddy named "Steven" have each been increased by 10 per cent because the user frequently calls these two individuals from home. In this example, "Irvin" would be dialed instead of "Mervyn" because "Irvin" now has the greatest confidence measure.

In other applications of the invention, the stored data corresponding to a stored pattern may be other than a telephone number. For example, the stored patterns may correspond to some other type of electronic address such as an e-mail address or Uniform Resource Locator (URL).

As an example of a different application of the invention, the aspect described in connection with Fig. 6 may be applied to a voice-driven e-mail client program running on mobile communication device 165. A voice-driven e-mail client allows a user to control its functions, at least in part, by speaking commands instead of using a keyboard or mouse, as is known in the art. For example, the user might be able to create a new

message by saying "new" and address the message by saying a label such as "Julie." As with voice-dialing, a problem arises if there are multiple people with the same label or if the speech recognizer assigns a nearly equal confidence measure to a pattern associated with a different label such as "Jerry." In a manner similar to that illustrated in Fig. 5, a record may be kept of the user-defined, named locations such as "Home" or "Office" from which past e-mail messages have been sent from mobile communication device 165. When an ambiguous situation occurs, the confidence measures associated with e-mail addresses frequently contacted from the current named location of mobile communication device 165 are modified to favor the corresponding e-mail addresses over those less frequently contacted from that same location. For example, the user may frequently e-mail "Julie," a neighborhood friend, from mobile communication device 165 while at home but almost never e-mail "Jerry," a network support technician at the user's workplace, from that location. If speech recognition circuit 110 produces nearly equal confidence measures for the patterns corresponding to "Julie" and "Jerry," the invention favors "Julie" based on the correlated location information contained in the record.

Yet another application for the invention is voice dictation, which allows a user to create or edit documents by speaking instead of typing. Based on the current location of mobile communication device 165, different recognition vocabularies may be favored. For example, if the user is at home, a particular household

vocabulary is more likely than when the user is at the office. Conversely, when the user is at the office, a technical or professional vocabulary is more likely than the household vocabulary. Therefore, the invention favors the appropriate vocabulary based on location context.

It will be appreciated by those of ordinary skill in the art that the present invention may be embodied in other specific forms without departing from the spirit or essential character thereof. The presently disclosed aspects are, therefore, considered in all respects to be illustrative and not restrictive.